



IFWO

RAW SEQUENCE LISTING

DATE: 08/31/2004

PATENT APPLICATION: US/10/782,257

TIME: 12:22:49

Input Set : N:\Crf3\RULE60\10782257.raw
 Output Set: N:\CRF4\08312004\J782257.raw

1 <110> APPLICANT: Biotica Technology Limited
 2 Leadlay, Peter F
 3 Pfizer, Inc.
 4 Staunton, James
 5 Cortes, Jesus
 6 McArthur, Hamish AI
 7 <120> TITLE OF INVENTION: Polyketides and their synthesis
 8 <130> FILE REFERENCE: IS/CP5787585
 9 <140> CURRENT APPLICATION NUMBER: US/10/782,257
 10 <141> CURRENT FILING DATE: 2004-02-19
 11 <150> PRIOR APPLICATION NUMBER: US/09/720,841
 12 <151> PRIOR FILING DATE: 2001-08-13
 13 <150> PRIOR APPLICATION NUMBER: GB 9814006.4
 14 <151> PRIOR FILING DATE: 1998-06-29
 15 <160> NUMBER OF SEQ ID NOS: 53
 16 <170> SOFTWARE: PatentIn Ver. 2.1
 18 <210> SEQ ID NO: 1
 19 <211> LENGTH: 398
 20 <212> TYPE: PRT
 21 <213> ORGANISM: Streptomyces sp. C5
 22 <400> SEQUENCE: 1
 Met Val Thr Gly Leu Gly Ile Val Ala Pro Asn Gly Leu Gly Val Gly
 1 5 10 15
 Ala Ile Trp Asp Ala Val Leu Asn Gly Arg Asn Gly Ile Gly Pro Leu
 20 25 30
 Arg Arg Phe Ala Asp Asp Gly Arg Leu Gly Arg Leu Ala Gly Glu Val
 35 40 45
 Ser Asp Phe Val Pro Glu Asp His Leu Pro Lys Arg Leu Leu Val Gln
 50 55 60
 Thr Asp Pro Met Thr Gln Met Thr Ala Leu Ala Ala Glu Trp Ala
 65 70 75 80
 Leu Arg Glu Ala Gly Cys Ala Pro Ser Ser Pro Leu Glu Ala Gly Val
 85 90 95
 Ile Thr Ala Ser Ala Ser Gly Gly Phe Ala Ser Gly Gln Arg Glu Leu
 100 105 110
 Gln Asn Leu Trp Ser Lys Gly Pro Ala His Val Ser Ala Tyr Met Ser
 115 120 125
 Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile Ala Ile Arg His
 130 135 140
 Asp Leu Arg Gly Pro Val Gly Val Val Ala Glu Gln Ala Gly Gly
 145 150 155 160
 Leu Asp Ala Leu Ala His Ala Arg Arg Lys Val Arg Gly Gly Ala Glu
 165 170 175

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/782,257

DATE: 08/31/2004

TIME: 12:22:49

Input Set : N:\Crf3\RULE60\10782257.raw
 Output Set: N:\CRF4\08312004\J782257.raw

45 Leu Ile Val Ser Gly Ala Met Asp Ser Ser Leu Cys Pro Tyr Gly Met
 46 180 185 190
 47 Ala Ala Gln Val Arg Ser Gly Arg Leu Ser Gly Ser Asp Asp Pro Thr
 48 195 200 205
 49 Ala Gly Tyr Leu Pro Phe Asp Arg Arg Ala Ala Gly His Val Pro Gly
 50 210 215 220
 51 Glu Gly Gly Ala Ile Leu Ala Val Glu Asp Ala Glu Arg Val Ala Glu
 52 225 230 235 240
 53 Arg Gly Gly Lys Val Tyr Gly Ser Ile Ala Gly Thr Ala Ser Phe Asp
 54 245 250 255
 55 Pro Pro Pro Gly Ser Gly Arg Pro Ser Ala Leu Ala Arg Ala Val Glu
 56 260 265 270
 57 Thr Ala Leu Ala Asp Ala Gly Leu Asp Arg Ser Asp Ile Ala Val Val
 58 275 280 285
 59 Phe Ala Asp Gly Ala Ala Val Gly Glu Leu Asp Val Ala Glu Ala Glu
 60 290 295 300
 61 Ala Leu Ala Ser Val Phe Gly Pro His Arg Val Pro Val Thr Val Pro
 62 305 310 315 320
 63 Lys Thr Leu Thr Gly Arg Leu Tyr Ser Gly Ala Gly Pro Leu Asp Val
 64 325 330 335
 65 Ala Thr Gly Leu Ala Leu Arg Asp Glu Val Val Pro Ala Thr Gly
 66 340 345 350
 67 His Val His Pro Asp Pro Asp Leu Pro Leu Asp Val Val Thr Gly Arg
 68 355 360 365
 69 Pro Arg Ala Met Ala Asp Ala Arg Ala Ala Leu Val Val Ala Arg Gly
 70 370 375 380
 71 His Gly Gly Phe Asn Ser Ala Leu Val Val Arg Gly Ala Ala
 72 385 390 395
 74 <210> SEQ ID NO: 2
 75 <211> LENGTH: 425
 76 <212> TYPE: PRT
 77 <213> ORGANISM: Streptomyces peucetius
 78 <400> SEQUENCE: 2
 79 Met Thr Gly Thr Ala Ala Arg Thr Ala Ser Ser Gln Leu His Ala Ser
 80 1 5 10 15
 81 Pro Ala Gly Arg Arg Gly Leu Arg Gly Arg Ala Val Val Thr Gly Leu
 82 20 25 30
 83 Gly Ile Val Ala Pro Asn Gly Leu Gly Val Gly Ala Tyr Trp Asp Ala
 84 35 40 45
 85 Val Leu Asn Gly Arg Asn Gly Ile Gly Pro Leu Arg Arg Phe Thr Gly
 86 50 55 60
 87 Asp Gly Arg Leu Gly Arg Leu Ala Gly Glu Val Ser Asp Phe Val Pro
 88 65 70 75 80
 89 Glu Asp His Leu Pro Lys Arg Leu Leu Ala Gln Thr Asp Pro Met Thr
 90 85 90 95
 91 Gln Tyr Ala Leu Ala Ala Glu Trp Ala Leu Arg Glu Ser Gly Cys
 92 100 105 110
 93 Ser Pro Ser Ser Pro Leu Glu Ala Gly Val Ile Thr Ala Ser Ala Ser
 94 115 120 125

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/782,257

DATE: 08/31/2004

TIME: 12:22:49

Input Set : N:\Crf3\RULE60\10782257.raw

Output Set: N:\CRF4\08312004\J782257.raw

95 Gly Gly Phe Ala Phe Gly Gln Arg Glu Leu Gln Asn Leu Trp Ser Lys
 96 130 135 140
 97 Gly Pro Ala His Val Ser Ala Tyr Met Ser Phe Ala Trp Phe Tyr Ala
 98 145 150 155 160
 99 Val Asn Thr Gly Gln Ile Ala Ile Arg His Asp Leu Arg Gly Pro Val
 100 165 170 175
 101 Gly Val Val Val Ala Glu Gln Ala Gly Gly Leu Asp Ala Leu Ala His
 102 180 185 190
 103 Ala Arg Arg Lys Val Arg Gly Gly Ala Glu Leu Ile Val Ser Gly Ala
 104 195 200 205
 105 Val Asp Ser Ser Leu Cys Pro Tyr Gly Met Ala Ala Gln Val Lys Ser
 106 210 215 220
 107 Gly Arg Leu Ser Gly Ser Asp Asn Pro Thr Ala Gly Tyr Leu Pro Phe
 108 225 230 235 240
 109 Asp Arg Arg Ala Ala Gly His Val Pro Gly Glu Gly Gly Ala Ile Leu
 110 245 250 255
 111 Thr Val Glu Asp Ala Glu Arg Ala Ala Glu Arg Gly Ala Lys Val Tyr
 112 260 265 270
 113 Gly Ser Ile Ala Gly Tyr Gly Ala Ser Phe Asp Pro Pro Pro Gly Ser
 114 275 280 285
 115 Gly Arg Pro Ser Ala Leu Ala Arg Ala Val Glu Thr Ala Leu Ala Asp
 116 290 295 300
 117 Ala Gly Leu Asp Gly Ser Asp Ile Ala Val Val Phe Ala Asp Gly Ala
 118 305 310 315 320
 119 Ala Val Pro Glu Leu Asp Ala Ala Glu Ala Glu Ala Leu Ala Ser Val
 120 325 330 335
 121 Phe Gly Pro Arg Arg Val Pro Val Thr Val Pro Lys Thr Leu Thr Gly
 122 340 345 350
 123 Arg Leu Tyr Ser Gly Ala Gly Pro Leu Asp Val Ala Thr Ala Leu Leu
 124 355 360 365
 125 Ala Leu Arg Asp Glu Val Val Pro Ala Thr Ala His Val Asp Pro Asp
 126 370 375 380
 127 Pro Asp Leu Pro Leu Asp Val Val Thr Gly Arg Pro Arg Ser Leu Ala
 128 385 390 395 400
 129 Asp Ala Arg Ala Ala Leu Leu Val Ala Arg Gly Tyr Gly Gly Phe Asn
 130 405 410 415
 131 Ser Ala Leu Val Val Arg Gly Ala Ala
 132 420 425
 134 <210> SEQ ID NO: 3
 135 <211> LENGTH: 407
 136 <212> TYPE: PRT
 137 <213> ORGANISM: Streptomyces coelicolor
 138 <400> SEQUENCE: 3
 139 Met Ser Val Leu Ile Thr Gly Val Val Ala Pro Asn Gly Leu
 140 1 5 10 15
 141 Gly Leu Ala Pro Tyr Trp Ser Ala Val Leu Asp Gly Arg His Gly Leu
 142 20 25 30
 143 Gly Pro Val Thr Arg Phe Asp Val Ser Arg Tyr Pro Ala Thr Leu Ala
 144 35 40 45

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/782,257

DATE: 08/31/2004
TIME: 12:22:49

Input Set : N:\Crf3\RULE60\10782257.raw
Output Set: N:\CRF4\08312004\J782257.raw

145 Gly Gln Ile Asp Asp Phe His Ala Pro Asp His Ile Pro Gly Arg Leu
146 50 55 60
147 Leu Pro Gln Thr Asp Pro Ser Thr Arg Leu Ala Leu Thr Ala Ala Asp
148 65 70 75 80
149 Trp Ala Leu Gln Asp Ala Lys Ala Asp Pro Glu Ser Leu Thr Asp Tyr
150 85 90 95
151 Asp Met Gly Val Val Thr Ala Asn Ala Cys Gly Gly Phe Asp Phe Thr
152 100 105 110
153 His Arg Glu Phe Arg Lys Leu Trp Ser Glu Gly Pro Lys Ser Val Ser
154 115 120 125
155 Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile
156 130 135 140
157 Ser Ile Arg His Gly Met Arg Gly Pro Ser Ser Ala Leu Val Ala Glu
158 145 150 155 160
159 Gln Ala Gly Gly Leu Asp Ala Leu Gly His Ala Arg Arg Thr Ile Arg
160 165 170 175
161 Arg Gly Thr Pro Leu Val Val Ser Gly Gly Val Asp Ser Ala Leu Asp
162 180 185 190
163 Pro Trp Gly Trp Val Ser Gln Ile Ala Ser Gly Arg Ile Ser Thr Ala
164 195 200 205
165 Thr Asp Pro Asp Arg Ala Tyr Leu Pro Phe Asp Glu Arg Ala Ala Gly
166 210 215 220
167 Tyr Val Pro Gly Glu Gly Gly Ala Ile Leu Val Leu Glu Asp Ser Ala
168 225 230 235 240
169 Ala Ala Glu Ala Arg Gly Arg His Asp Ala Tyr Gly Glu Leu Ala Gly
170 245 250 255
171 Cys Ala Ser Thr Phe Asp Pro Ala Pro Gly Ser Gly Arg Pro Ala Gly
172 260 265 270
173 Leu Glu Arg Ala Ile Arg Leu Ala Leu Asn Asp Ala Gly Thr Gly Pro
174 275 280 285
175 Glu Asp Val Asp Val Val Phe Ala Asp Gly Ala Gly Val Pro Glu Leu
176 290 295 300
177 Asp Ala Ala Glu Ala Arg Ala Ile Gly Arg Val Phe Gly Arg Glu Gly
178 305 310 315 320
179 Val Pro Val Thr Val Pro Lys Thr Thr Thr Gly Arg Leu Tyr Ser Gly
180 325 330 335
181 Gly Gly Pro Leu Asp Val Val Thr Ala Leu Met Ser Leu Arg Glu Gly
182 340 345 350
183 Val Ile Ala Pro Thr Ala Gly Val Thr Ser Val Pro Arg Glu Tyr Gly
184 355 360 365
185 Ile Asp Leu Val Leu Gly Glu Pro Arg Ser Thr Ala Pro Arg Thr Ala
186 370 375 380
187 Leu Val Leu Ala Arg Gly Arg Trp Gly Phe Asn Ser Ala Ala Val Leu
188 385 390 395 400
189 Arg Arg Phe Ala Pro Thr Pro
190 405
192 <210> SEQ ID NO: 4
193 <211> LENGTH: 403
194 <212> TYPE: PRT

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/782,257

DATE: 08/31/2004
TIME: 12:22:49

Input Set : N:\Crf3\RULE60\10782257.raw
Output Set: N:\CRF4\08312004\J782257.raw

195 <213> ORGANISM: Saccharopolyspora hirsuta
196 <400> SEQUENCE: 4
197 Met Ser Thr Trp Val Thr Gly Met Gly Val Val Ala Pro Asn Gly Leu
198 1 5 10 15
199 Gly Ala Asp Asp His Trp Ala Ala Thr Leu Lys Gly Arg His Gly Ile
200 20 25 30
201 Ser Arg Leu Ser Arg Phe Asp Pro Thr Gly Tyr Pro Ala Glu Leu Ala
202 35 40 45
203 Gly Gln Val Leu Asp Phe Asp Ala Thr Glu His Leu Pro Lys Arg Leu
204 50 55 60
205 Leu Pro Gln Thr Asp Val Ser Thr Arg Phe Ala Leu Ala Ala Ala Ala
206 65 70 75 80
207 Trp Ala Leu Ala Asp Ala Glu Val Asp Pro Ala Glu Leu Pro Glu Tyr
208 85 90 95
209 Gly Thr Gly Val Ile Thr Ser Asn Ala Thr Gly Gly Phe Glu Phe Thr
210 100 105 110
211 His Arg Glu Phe Arg Lys Leu Trp Ala Gln Gly Pro Glu Phe Val Ser
212 115 120 125
213 Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile
214 130 135 140
215 Ser Ile Arg His Gly Leu Arg Gly Pro Gly Ser Val Leu Val Ala Glu
216 145 150 155 160
217 Gln Ala Gly Gly Leu Asp Ala Val Gly His Gly Gly Ala Val Arg Asn
218 165 170 175
219 Gly Thr Pro Met Val Val Thr Gly Gly Val Asp Ser Ser Phe Asp Pro
220 180 185 190
221 Trp Gly Trp Val Ser His Val Ser Ser Gly Arg Val Ser Arg Ala Thr
222 195 200 205
223 Asp Pro Gly Arg Ala Tyr Leu Pro Phe Asp Val Ala Ala Asn Gly Tyr
224 210 215 220
225 Val Pro Gly Glu Gly Gly Ala Ile Leu Leu Leu Glu Asp Ala Glu Ser
226 225 230 235 240
227 Ala Lys Ala Arg Gly Ala Thr Gly Tyr Gly Glu Ile Ala Gly Tyr Ala
228 245 250 255
229 Ala Thr Phe Asp Pro Ala Pro Gly Ser Glu Arg Pro Pro Ala Leu Arg
230 260 265 270
231 Arg Ala Ile Glu Leu Ala Leu Ala Asp Ala Glu Leu Arg Pro Glu Gln
232 275 280 285
233 Val Asp Val Val Phe Ala Asp Ala Ala Gly Val Ala Glu Leu Asp Ala
234 290 295 300
235 Ile Glu Ala Ala Ala Ile Arg Glu Leu Phe Gly Pro Ser Gly Val Pro
236 305 310 315 320
237 Val Thr Ala Pro Lys Thr Met Thr Gly Arg Leu Tyr Ser Gly Gly
238 325 330 335
239 Pro Leu Asp Leu Val Ala Ala Leu Leu Ala Ile Arg Asp Gly Val Ile
240 340 345 350
241 Pro Pro Thr Val His Thr Ala Glu Pro Val Pro Glu His Gln Leu Asp
242 355 360 365
243 Leu Val Thr Gly Asp Pro Arg His Gln Gln Leu Gly Thr Ala Leu Val

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/782,257

DATE: 08/31/2004

TIME: 12:22:50

Input Set : N:\Crf3\RULE60\10782257.raw

Output Set: N:\CRF4\08312004\J782257.raw